



# Spatial validation of two simulation models: the example of timber plantations in Chile

Nicolas Maestriperier, Martin Paegelow

## ► To cite this version:

Nicolas Maestriperier, Martin Paegelow. Spatial validation of two simulation models: the example of timber plantations in Chile. *Cybergeog: Revue européenne de géographie / European journal of geography*, 2013, Systèmes, modélisation, géostatistiques, pp.en ligne. hal-01092758

**HAL Id: hal-01092758**

**<https://hal.science/hal-01092758>**

Submitted on 9 Dec 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Validation spatiale de deux modèles de simulation : l'exemple des plantations industrielles au Chili

### *Spatial validation of two simulation models: the example of timber plantations in Chile*

#### Résumé

La commune de San Juan de la Costa (province d'Osorno – 10<sup>ème</sup> Région de Los Lagos) connaît depuis la fin des années 1980 une expansion remarquable des plantations de pins et d'eucalyptus. Ces boisements sont à l'origine de nombreuses perturbations environnementales mais aussi sociales dans une région considérée comme un *hotspot* de biodiversité. Si les processus de changement d'occupation et d'utilisation du sol (LUCC) sont dans l'ensemble bien connus, les zones de forêts natives susceptibles d'être converties en plantations industrielles sont quant à elles inconnues. Cet article se focalise sur la validation spatiale de deux modèles de simulation afin de proposer des interprétations prospectives en se basant sur une approche supervisée (chaînes de Markov couplées à une évaluation multicritère) et une approche automatique (réseau neuronal multicouches). Les deux modèles intègrent la connaissance des dynamiques passées grâce à l'analyse diachronique de trois images satellites LANDSAT (1986-1999-2008).

#### Abstract

Since the end of the 1980s, the commune of San Juan de la Costa (province of Osorno – 10th region of Los Lagos), is experiencing a remarkable expansion of pine and eucalyptus plantations. These afforestation are at the roots of many environmental and social disturbances in an area considered as a hotspot for biodiversity. If the Land Use and Land Cover Change (LUCC) processes are widely known, the native forests areas likely to be converted into industrial plantations remain unknown. The following article focuses on the spatial validation of two simulation models in order to suggest prospective interpretations based on a supervised approach (Markov chain coupled with a multi-criteria evaluation) and an automatic approach (multilayer neural network). Both models build in the knowledge of past dynamics thanks to the diachronic analysis of three satellite images LANDSAT (1986-1999-2008).

**Mots-clés:** Modèles spatialement explicites, Plantations industrielles, CA-Markov, Land Change Modeler, Validation, Prospective

**Keywords:** Spatially explicit models, Industrial plantations, CA-Markov, Land Change Modeler, Validation, Prospective

## Introduction

"La transformation des terres englobe une grande variété d'activités qui varient considérablement dans leur intensité et leurs conséquences" (Vitousek, 1997). Le développement remarquable de la sylviculture intensive à travers le monde depuis les années 1960 (Bull, 2006) constitue même une des manifestations les plus répandues du changement global (Arroyo et al., 2000).

La demande croissante en fibres par les industries de pâtes à papier (Evans et Turnbull, 2004) et les politiques d'incitations au développement des plantations industrielles n'ont fait que renforcer ce processus. Tel est le cas du centre-sud chilien où se produit, depuis l'application du Décret-Loi 701<sup>1</sup> en 1974, une rapide expansion des monocultures de *Pinus radiata* et d'*Eucalyptus*. Entre 1975 et 2007 la superficie forestière moyenne annuelle plantée (boisement et reboisement) à l'échelle nationale était de 95 141 ha, pour atteindre une superficie de 2,2M d'ha en 2007 (INFOR, 2008).

Malgré la reconnaissance de la durabilité des activités sylvicoles chiliennes par l'OCDE en 2005 (Arana Silva, 2009), la communauté scientifique s'accorde sur les limites environnementales et socio-économiques engendrées par de telles pratiques. Aux problèmes de pauvreté et d'exode ruraux et d'expulsion de minorité ethnique mapuche (Lara, 1985 ; Leyton Vasquez, 2009), s'ajoutent la perte de biodiversité (Sheldon, 1989 in Winjum, 1997 ; Donoso et Otero, 2005), la perturbation des cycles hydrologiques et l'acidification des sols (Cannell, 1999). En outre, la substitution des forêts natives par des monocultures a été reconnue comme l'une des principales menace observée dans cette zone (Lara, 1996 ; Cavelier et Tecklin, 2005 ; Donoso et Otero, 2005 ;

<sup>1</sup> Ce décret incite au boisement sur des terrains d'aptitudes préférentiellement forestières, dépourvus de couverture arborée et dans la plupart des cas fortement érodés. Il permet de subventionner entre 75 et 90% des coûts de plantation et d'exempter d'impôts les plantations effectuées sur des terrains classés d'aptitude forestière (Barrué Pastor, 2004).



Echeverria et al., 2006 ; Altamirano et Lara, 2010). L'enjeu est effectivement de taille car la majorité de cette exploitation forestière se situe au sein même de l'Ecorégion des Forêts Tempérées Humides Valdivienne qui fait partie des 238 Ecorégions identifiées par le WWF au niveau mondial. Elle fait l'objet d'une des plus fortes priorités de conservation et de restauration dans le monde et fait partie des 25 *hotspots* pour la conservation de la biodiversité (Myers, 2000).

Les objectifs de cette étude sont : (1) de calibrer et de valider deux modèles spatialement explicites (CA-Markov et LCM) ; (2) de comparer les résultats de deux modèles et de mettre en exergue les espaces à enjeux (plantations exotiques). Les résultats obtenus permettront de réaliser (hors cadre de cet article) une simulation en 2017<sup>2</sup> et serviront de support aux décideurs dans l'optique d'une meilleure gestion des ressources forestières. Ainsi, après avoir présenté l'aire d'étude et les données mobilisées, les deux modèles ainsi que les méthodes de validation seront détaillés. Enfin l'article se focalisera sur la qualité des résultats obtenus et discutera de l'intérêt et des limites de ces modèles.

## Aire d'étude et données mobilisées

### Aire d'étude

La commune de San Juan de la Costa (40°14-40°44S de latitude et 73°18-73°48W de longitude) est l'une des sept communes de la province d'Osorno (10<sup>ème</sup> Région des Lacs) et couvre une superficie approximative de 154000 hectares (figure 1). La zone réunit deux des paysages caractéristiques du centre-sud chilien<sup>3</sup>. La précordillère côtière (à l'est) avec une altitude moyenne de 300 mètres, se caractérise par de petites propriétés tournées pour la plupart d'entre-elles vers des activités agropastorales, et quelques grandes exploitations de plantations industrielles exotiques aux mains d'entreprises forestières. Les forêts natives s'y trouvent fortement fragmentées (Saez, 2002 ; Sánchez Álvarez, 2004). La Cordillère de la Côte (à l'ouest) culmine à environ 800m au-dessus du niveau de mer et reçoit d'abondantes précipitations (3500mm/an) (Muñoz, 2004). Elle abrite les forêts valdiviennes de *Nothofagus* et d'alerce (*Fitzroya cupressoides*). Le fort taux d'endémisme s'explique par la "non glaciation" de la Cordillère côtière durant le Pléistocène, devenant un refuge pour de nombreuses espèces végétales et animales (Villagran, 1991).

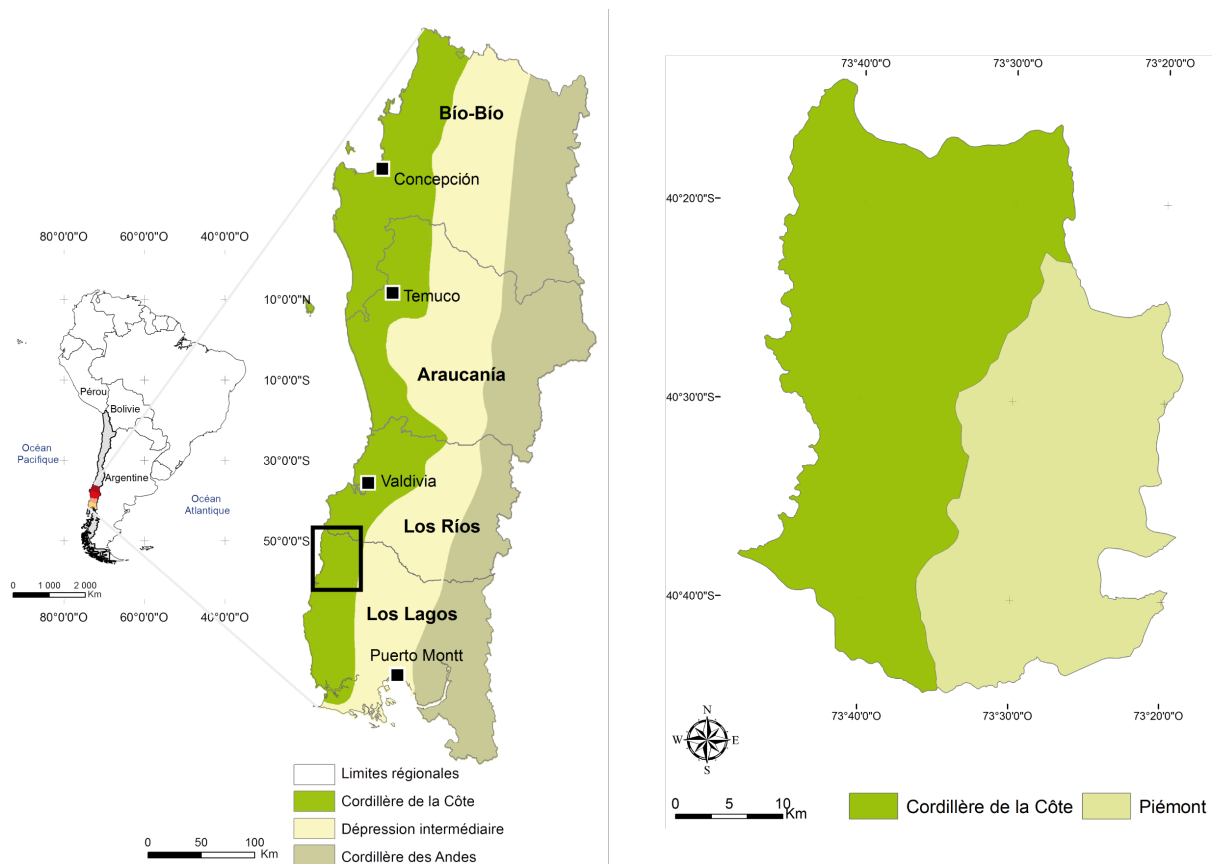
D'après le dernier recensement de la population (INE, 2002) sa population était de 8831 habitants avec une forte part de ruraux (7929 habitants soit 90%) et avec une majorité ethnique mapuche-huilliche (60% de la population totale). Sur les 62 communautés indigènes recensées en 2000 (Barrué-Pastor, 2004) 44 d'entre-elles se situaient sur la précordillère côtière. La commune connaît un taux de pauvreté très élevé. Selon l'Encuesta de Caracterización Socioeconómica Nacional (CASEN) publiée par le Ministerio de Planificación, le taux de pauvres (indigents et non indigents) s'élevait à 32,4% de la population communale en 2000, 42,1% en 2003, pour atteindre les 14,3% en 2006, alors qu'à l'échelle du pays ce taux était respectivement de 20,2%, 18,7% et 13,7% (BCN, 2008).

---

<sup>2</sup> Le calibrage du modèle se base sur la période 1986-1999 pour simuler en 2008. La simulation atteindra 2017 (2008 + 9) afin de respecter le pas de temps de 9 ans compris entre 1999 et 2008.

<sup>3</sup> D'est en ouest on distingue : la Cordillère des Andes, la dépression intermédiaire (vallée centrale), la précordillère côtière (ou piémont) et le Cordillère de la Côte.

Figure 1 : Localisation de l'aire d'étude



## Données

Pour analyser les dynamiques d'utilisation et d'occupation du sol sur un gradient temporel satisfaisant (22 ans), nous avons acquis trois images satellites LANDSAT : 1986 (TM), 1999 (TM) et 2008 (TM) (path 233, row 88). Chaque image dont la résolution du pixel est de 30 m x 30 m, a été corrigée géométriquement et topographiquement. Une classification supervisée par maximum de vraisemblance a été réalisée en s'appuyant sur des relevés de terrains (300 points GPS) combinés à des photographies aériennes et des images satellites Google Earth, ainsi que sur les données digitalisées (actualisée en 2006) issues du Cadastre de la Corporación Nacional Forestal (CONAF et al., 1999). L'imagerie satellitale LANDSAT représente un outil puissant pour le suivi des dynamiques environnementales (Wulder et al., 2008). La technique par maximum de vraisemblance a été largement employée pour le suivi des changements d'occupation des sols (Schulz et al., 2010 ; Wyman et Stein, 2010) et l'étude de la fragmentation forestière (Echeverria et al., 2006 ; Echeverria et al., 2008). La précision globale de chaque classification a été de 84,3% pour 1986, 91,2% pour 1999 et de 87,5% pour 2008.

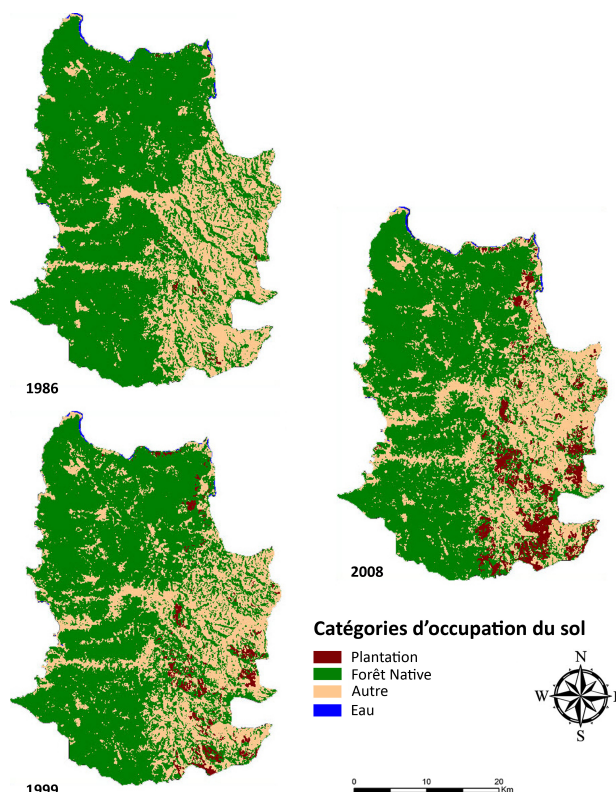
D'autres données digitalisées ont enfin été obtenues à savoir : (i) une carte des pentes, générée grâce à un Modèle Numérique de Terrain lui-même créé par des courbes de niveaux (pas de 50m) ainsi qu'une carte du réseau hydrographique et du réseau routier (2003) acquises auprès de l'Instituto Geográfico Militar (IGM) en 2008 et (ii) une carte d'altitude (CONAF, 2006).

La typologie de base comportait quinze classes afin d'appréhender les types d'occupation du sol avec une plus grande précision. Les confusions radiométriques (coupe rase et jeune plantation d'eucalyptus par exemple) couplées à la redondance de certaines classes nous amenèrent à les agréger en un nombre plus restreint. Six catégories d'occupation du sol ont ainsi été identifiées pour chaque image : (i) plantations (pins et eucalyptus adulte), (ii) forêt native adulte (FN), (iii) renoval et matorral -R&M- (forêt secondaire et fourré), (iv) prairie (pâturage, prairie de fauche dense et terre agricole), (v) sol nu (coupe rase, plantation jeune, prairie sèche, alerzales), et (vi) corps d'eau. Dans un dernier temps et afin de simplifier le processus de simulation prospective et de mettre en exergue la dynamique des plantations industrielles, la légende a été réduite à quatre types

d'occupation du sol : (i) plantations, (ii) forêt native, (iii) autre (R&M, prairie et sol nu), et (iv) corps d'eau (figure 2).

Avec plus de 96 000 ha en 2008, la forêt native couvre la quasi-totalité de la Cordillère de la Côte et est fragmentée sur le Piémont. Depuis 1986 sa superficie diminua de 13%. Les plantations industrielles se situent exclusivement sur le Piémont, et aux abords du fleuve Bueno et du Lac Rupanco au nord et nord-est de la commune. A défaut d'avoir occupé la plus faible superficie en 1986 avec 387 ha, les plantations ont connu la plus forte augmentation entre 1986 et 2008 avec plus de 2 455% (9 888 ha). La catégorie "autre" se répartie sur l'ensemble de la commune, majoritairement sur le Piémont et de façon plus éparse sur la Cordillère. Sa superficie passa de 42 907 ha en 1986 à plus de 47 000 ha 22 ans plus tard, soit une évolution de 9,5%.

Figure 2 : Visualisation diachronique de l'occupation du sol (1986-1999-2008) de l'occupation du sol



## Méthode

## L'intérêt de la prospective

Afin de répondre aux enjeux du développement soutenable, la nécessité d'entrevoir les changements futurs grâce à l'approche prospective devient incontournable (Hubert-Moy, 2006 ; Houet et al., 2008). La prospective en France remonte à Gaston Berger par la création en 1957 du Centre International de prospective et du World Futures Studies Federation par Bertrand de Jouvenel dans les années 1960. La diffusion de cette "philosophie et attitude" (Condorbes et Durance, 2004) dans le secteur public français se fera par la Délégation interministérielle à l'Aménagement du Territoire et à l'Attractivité Régionale (DATAR) par l'intermédiaire de Pierre Massé en publiant les "3 scénarios de l'inacceptable". L'histoire de la prospective et sa démarche méthodologique étant abondamment traitée, nous orientons le lecteur sur quelques références clés (Godet, 1986 ; De Jouvenel, 1999 ; Houet, 2006 ; Voiron-Canicio, 2006). Ces auteurs s'accordent sur le fait que cette démarche ne vise pas à prédire un futur. La prospective, propose cependant des ouvertures, des pistes sur l'image de ce *demain*. Il est question de construire, grâce à l'analyse du passé et des dynamiques actuelles, un socle d'hypothèses pertinentes considérées par Fabrice Hatem comme "*un regard sur l'avenir destiné à éclairer l'action présente*" (1993, p.18).

Ainsi, l'analyse du passé et du présent n'est pas la seule condition à l'explication du futur, dans la mesure où le futur est imprimé dans le présent (Godet, 1986) par l'anticipation qu'exerce un individu sur son environnement. La démarche participative représentée par la volonté des acteurs sociaux devient la clé de la prospective (Medina Vásquez et Ortégón, 2006).

## L'intérêt de l'approche spatiale

A l'instar de la nature autonome des études spatialement explicites et des approches socio-économiques pour appréhender les relations société/environnement (Veldkamp et Verburg, 2004), l'approche spatiale et la prospective possèdent encore de faibles connections. L'intérêt de la modélisation spatiale réside dans sa capacité à révéler les principales structures spatiales et leur dynamique et de favoriser la compréhension et l'explication de leurs causes et conséquences (Verburg et al., 2004). Veldkamp et Lambin (2001) insistent sur l'importance de la modélisation spatialement explicite en tant que technique pour la projection de voies alternatives vers le futur, alors que Paegelow et al. (2004) voient dans la modélisation prospective des données géographiques un défi majeur pour la recherche en géomatique.

Contrairement aux nombreuses études menées au sein du Projet Millénaire (Glenn et al., 2010), qui mettent en exergue la représentation mentale des scénarii dans une perspective humaine et sociale, la spatialisation de ces scénarii est fondamentale dans toute démarche prospective en géographie (Géoprospective) (Voinov-Canicio, 2006 ; Houet, 2006). Elle répond à ce que Medina Vásquez et Ortégón (2006) appellent la "construction sociale des futurs" qui se caractérise par (i) la détection des changements qui fait émerger une *image* du futur, (ii) la détermination de ce qui peut ou pas changer (*vision*), pour aboutir à l'identification des choses qui doivent changer (le *projet*).

Le recours à une approche spatiale contribue ainsi à une meilleure compréhension des changements d'utilisation et d'occupation des sols permettant a posteriori (i) d'identifier des espaces à enjeux pour la gestion et la conservation de la biodiversité ainsi que de (ii) localiser des conflits socio-spatiaux (Costanza et Voinov, 2004 ; Houet et al., 2008 ; Morán-Ordóñez et al., 2011).

## Une typologie des modèles LUCC

La modélisation des changements d'occupation et d'utilisation des sols ne se limite pas à une seule catégorie de modèle, ni à un seul domaine de recherche. Nous présenterons ici les modèles communément employés.

Tout d'abord les modèles stochastiques. Il s'agit de modèles probabilistes à événements discrets, c'est-à-dire qui n'évoluent pas de façon continue au cours du temps, mais avec des pas temporels discrets (Coquillard et Hill, 1997). Les chaînes de Markov font partie des méthodes les plus communément employées afin de prédire des changements entre plusieurs états de catégories (Pontius et Malanson, 2005). Elles ont été largement utilisées afin d'étudier les dynamiques de végétation (Coquillard et Hill, 1997 ; Paegelow et al., 2004 ; Paegelow et Camacho Olmedo, 2005 ; Camacho Olmedo et al., 2007). Selon Heiko Baltzer (2000, p.151) ces modèles ont la capacité de regrouper des informations très complexes permettant d'analyser des écosystèmes même si leurs processus sous-jacents ne sont pas totalement compris.

Les Automates cellulaires (AC) sont une autre forme de modèles stochastiques qui discrétisent un espace continu en une série de cellules. La littérature scientifique faisant usage des automates cellulaires est abondante, en témoigne la compilation réalisée par Baltzer et al. (1998). L'apport des AC dans l'analyse des processus écologiques réside dans la considération des interrelations spatiales entre les communautés végétales et les effets de voisinage qui les caractérisent.

Parmi les modèles à base d'intelligence artificielle, les modèles multi-agents (ou Systèmes Multi-Agents – SMA) (Bonnetoy et al., 2001 ; Parker et al., 2003 ; Daudé, 2005 ; Ferber, 2006) permettent d'appréhender et de représenter des systèmes complexes et dynamiques en modélisant directement les individus, leurs comportements et leurs interactions. Les SMA sont très largement employés par l'équipe du CIRAD à Montpellier notamment dans le cadre de la modélisation d'accompagnement (Collectif ComMod, 2005).

Les réseaux neuronaux artificiels (ANNS) font leur première apparition en tant que méthode permettant de modéliser des systèmes écologiques complexes dans l'étude de Colosanti (1991). Il existe plusieurs types de réseaux neuronaux (Lek et Guégan, 1999). Notre attention se porte sur le perceptron multi-couche (ou multi-layer feed-forward - MLP) car c'est l'une des architectures les plus employées pour l'étude des dynamiques d'occupations et d'utilisations du sol (Lek et Guégan, 1999 ; Mas et al. 2004 ; Paegelow et al., 2004). A l'image des modèles stochastiques, le MLP permet d'examiner et de modéliser des relations complexes entre plusieurs variables. Sa capacité à appréhender des relations non-linéaires en fait un outil puissant pour l'analyse des

processus écologiques. En terme de logiciel de modélisation utilisant le MLP, Land Change Modeler (Eastman, 2006 ; Oñate-Valdivieso et Bosque Sendra, 2010) et DINAMICA EGO (Soares-Filho et al., 2009) sont les plus utilisés.

Les modèles empiriques-statistiques spatialement explicites sont largement employés pour prédire les dynamiques de la végétation notamment dans le domaine des LUCC (Perry et Millington, 2008). Le modèle CLUE (Conversion of Land Use and its Effects) a été développé pour prédire et visualiser les dynamiques d'occupation du sol en se basant sur les relations empiriquement quantifiées entre ces dynamiques et leurs facteurs explicatifs (biophysiques et humains) (Veldkamp et Fresco, 1996 ; Verburg et al., 2002). A la différence de CA-Markov, Land Change Modeler et DINAMICA EGO qui utilisent les chaînes de Markov pour estimer les quantités de changement, le modélisateur doit intégrer les quantités (en pixels) de changements pour chaque année de simulation (Mas et al., sous presse).

## Deux modèles : CA-Markov et Land Change Modeler (LCM)

Afin de modéliser la dynamique des plantations industrielles et de simuler (dans le cadre d'un scénario tendanciel d'évolution) l'occupation du sol à l'horizon de 2017, il convient au préalable de valider le modèle sur des données connues. L'image de 2008 =  $t_2$  étant la plus récente, elle fera l'objet d'une première simulation-test, calibrée par deux dates antérieures ( $t_0 = 1986$  et  $t_1 = 1999$ ) (chaîne de Markov d'ordre 2). Les images de 1986 et 1999 servent de base pour extrapoler les quantités de la future occupation du sol. Il s'agit d'une extrapolation linéaire car la simulation se base sur deux points dans le temps afin de calibrer le modèle. La comparaison entre les simulations et l'observation 2008 permettra, d'une part, de quantifier les erreurs des deux modèles, et, d'autre part, de localiser ces erreurs à l'échelle de la commune. Les deux modèles présentés sont implémentés sous le logiciel de géostatistique IDRISI 15.0 (Eastman, 2006).

### CA-Markov

La CA-Markov est une procédure de prédiction LUCC combinant chaînes de Markov (quantité), évaluation multicritère (EMC) (localisation) et filtrage, une procédure qualifiée par les éditeurs du logiciel d'automate cellulaire (AC) (Eastman, 2006).

Les chaînes markoviennes analysent deux images d'occupation du sol à des dates différentes et produisent deux matrices de transition (probabilité et surface concernée en pixels pour la persistance et la transition), et un ensemble d'images de probabilité conditionnelle (Eastman, 2006). Elles permettent *"de calculer un état futur à partir de l'état présent connu, en se basant sur l'observation des évolutions passées et leur probabilité"* (Antoni, 2006). Les limites de l'analyse markovienne caractérisées par la non-allocation spatiale de ces transitions, peuvent être corrigées en ayant recours à une évaluation multicritère (Paegelow et Camacho Olmedo, 2005).

L'objectif de l'évaluation multicritère est de construire des cartes d'aptitude ou de probabilité en intégrant un ensemble de critères mesurables et cartographiables. Ces cartes pouvant ainsi servir à l'élaboration de choix stratégiques d'intervention. Le modélisateur contrôle le processus de modélisation en identifiant et en caractérisant les facteurs explicatifs grâce à une approche experte. Dans le cadre de l'évaluation multicritère on distingue deux types de critères : (i) les facteurs et (ii) les contraintes. Un facteur est un critère qui renforce ou réduit la pertinence d'une variable particulière pour l'activité étudiée. Il indique le degré d'aptitude de chaque unité de l'espace, et est mesuré sur une échelle continue allant de 0 à 255. Les variables ayant des entités différentes, cette échelle d'indice permet ainsi de standardiser l'ensemble des facteurs, les rendant comparables. La phase de standardisation se fait par une fonction d'appartenance de logique floue. Une contrainte limite, quant à elle, l'espace en considération dans la mesure où elle masque les zones inaptées aux objectifs fixés. C'est une variable booléenne, les zones exclues de l'analyse seront codées 0, et celles prises en compte seront codées 1. S'en suit une phase de pondération des facteurs basées sur la technique de comparaison par paires (*pairwise comparison*) dans le contexte de processus décisionnel appelé Analytical hierarchy process (AHP). Ainsi, les facteurs sont comparés deux par deux dans une matrice de comparaison dite de Saaty (Saaty, 1977) et ce en fonction de leur importance relative par rapport à l'objectif fixé. Enfin, une deuxième série de poids (poids d'ordre) est appliquée aux facteurs dans le but de contrôler leur niveau de compensation ainsi que le niveau de risque dans la détermination des aptitudes.

Les cartes d'aptitudes issues de l'EMC sont combinées aux cartes de probabilités de transition dérivant du calcul des chaînes de Markov. A cela s'ajoute l'application d'un filtre de contiguïté (5x5) qui a pour but d'éliminer les occurrences isolées et de renforcer les zones d'agglomérations afin d'éviter l'effet "pixélisé" (*pepper and salt*). Le traitement tient compte des quantités markoviennes ce qui explique un fonctionnement en boucle.

## Land Change Modeler

Ce modèle a été développé afin d'étudier les LUCC, de projeter leur dynamique dans le futur et d'en évaluer l'impact sur la biodiversité (Eastman, 2006). Une première estimation des changements (pertes, gains, changements nets, constance) est réalisée par la création de graphiques et de cartes de changements. Cette phase d'analyse permet la mise en place de sous-modèles de transition potentielle entre plusieurs catégories d'occupation du sol tout en tenant compte des variables statiques ou dynamiques. Les variables qui sont intégrées à chaque sous-modèle sont préalablement évaluées avec le coefficient V de Cramer afin de connaître leur pouvoir explicatif. Pour modéliser les changements d'occupation et d'utilisation des sols, LCM utilise soit la technique des réseaux neuronaux artificiels multicouches (*Multi-Layer Perceptron* ou MLP), soit une régression logistique. Nous avons privilégié le réseau neuronal car il est plus performant que le modèle de régression multiple, notamment dans les systèmes complexes et non-linéaires (Karul et Soyupak, 2006), mais peut conduire à "un sur-ajustement de la relation probabilité de changement / variables explicatives et du poids respectifs des variables aux caractéristiques de la période de calibrage" (Mas *et al.*, sous presse, p.13). Par ailleurs le MLP est vivement recommandé par les éditeurs (Eastman, 2009).

Un réseau neuronal de type MLP est constitué d'unités (ou nœuds) interconnectées de manière hiérarchique, avec une couche d'entrée (*input layer*), une ou plusieurs couches cachées (*hidden layer*) qui joue le rôle d'une boîte noire, et la couche de sortie (*output layer*). L'architecture de ce réseau et son fonctionnement sont présentés en détail par Paegelow *et al.* (2004)<sup>4</sup>.

La procédure d'allocation spatiale des changements d'utilisation et d'occupation du sol se fait par affectation multi-objectif (MOLA) qui se couple à l'estimation de la quantité de changement effectuée par des chaînes de Markov. Le MOLA prend en compte la totalité des conversions afin de lister les catégories LUC perdant en surface et celles en gagnant sur les premières. Le résultat est une carte de transition potentielle pour chaque transition (Eastman, 2009).

## Détermination des facteurs explicatifs

Le nombre de variables explicatives qui seront intégrées aux modèles est contraint par leur disponibilité, leur spatialisation ainsi que par leur influence sur la localisation et les changements des types d'occupation du sol. Nous sommes conscients que le nombre de facteurs présentés et intégrés est restreint comparativement à la palette des variables potentiellement explicatives. La légitimité de ces données pour calibrer le modèle doit être discutée. En effet certaines d'entre-elles sont postérieures à 1999 ce qui fausse le calibrage et ainsi l'extrapolation des dynamiques d'occupations du sol. Les données sur le réseau routier datent de 2003 et devraient donc être exclues, de même que les zones urbaines qui datent de 2006. Cependant, et après vérification auprès d'experts, les axes de communications sont restés relativement stables entre 1999 et 2003. Seuls quelques chemins ont été construits. Enfin nous considérons que les zones urbaines sont restées stables en 7 ans et que leur expansion dans une commune rurale est minime. Les variables utilisées dans les deux modèles sont identiques hormis les deux contraintes utilisées lors de l'évaluation multicritère.

La procédure d'EMC mobilise huit critères spatialement explicites répartis en trois catégories (figure 3)<sup>5</sup> : (1) les facteurs socio-économiques, (2) les facteurs environnementaux et (3) les contraintes. Six facteurs sont analysés avec deux facteurs relevant de la socio-économie : (i) la distance aux axes de communication (DAC) et (ii) la distance aux zones urbaines (DZU). Nous considérons que ces facteurs anthropiques ont une influence directe à la fois sur la motivation des propriétaires car elles déterminent le coût final du produit dans le cas de plantations, et sur le processus de déforestation. La distance joue un rôle clé sur l'ensemble des interactions spatiales de même que sur la localisation des entités géographiques.

Quatre autres facteurs relevant de l'environnement: (iii) l'occupation du sol en 1999 (OS) qui est une variable dépendante, (iv)<sup>6</sup> la distance aux entités de mêmes natures (DEN), (v) l'altitude (ALT) et (vi) la pente (PTE). Enfin

<sup>4</sup> Nous renvoyons le lecteur sur les travaux de Lek et Guégan (1999), Mas *et al.* (2004), et Parizeau (2004) pour une description détaillée de sa structure et de son fonctionnement.

<sup>5</sup> Chaque variable (i à vi) ayant une unité propre (mètres, degrés, etc.) celles-ci ont été recodées ou standardisées sur une échelle commune d'aptitude comprise entre 0 et 255 afin de les rendre comparables.

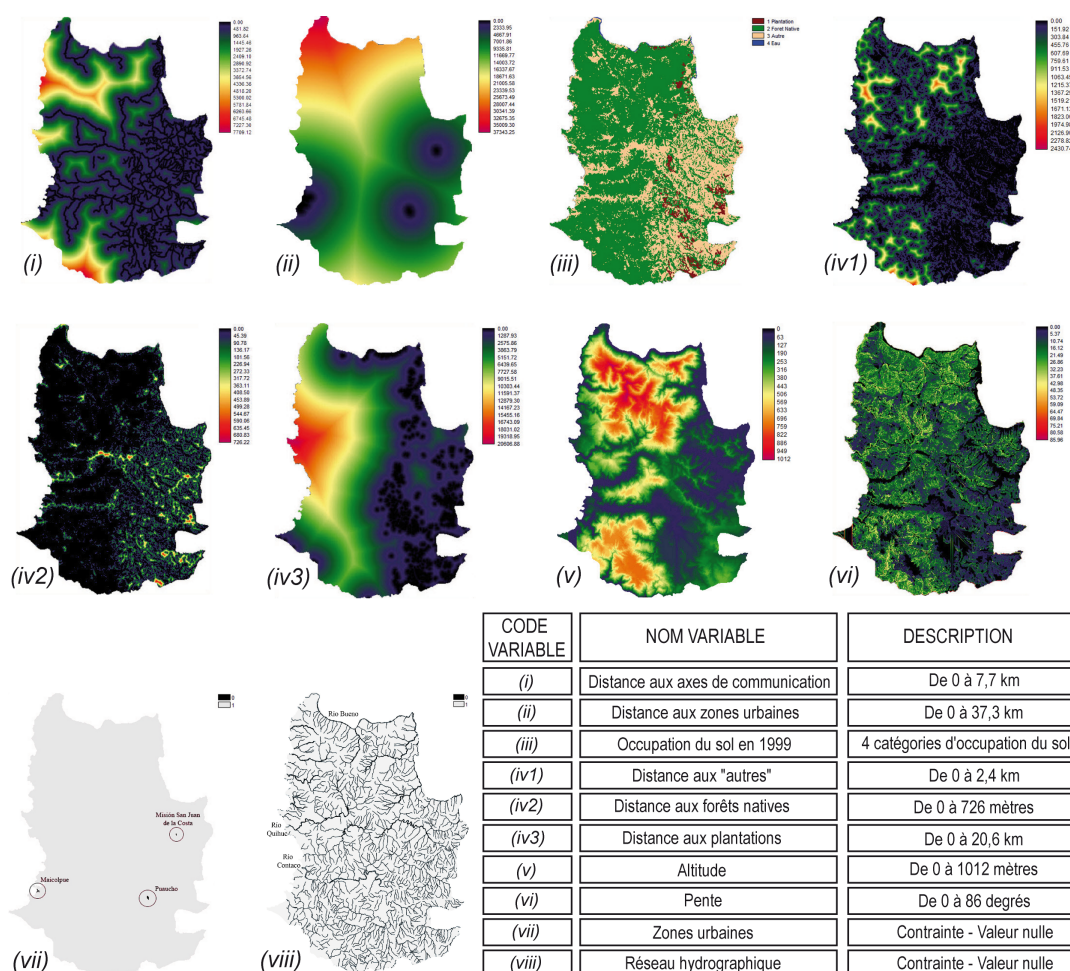
<sup>6</sup> iv1: proximité à "autre" 1986; iv2: proximité à la forêt native 1986; iv3: proximité aux plantations 1986. L'hypothèse de base est que le voisinage immédiat d'un type d'occupation du sol influence l'apparition d'entités de même nature. Les nouveaux espaces de forêt native correspondent potentiellement à des forêts natives régénérées ou à des plantations de forêt

deux critères contraignants sont intégrés et évalués dans le modèle : (vii) les zones urbaines (CZU) et (viii) le réseau hydrographique (CRH). Enfin, une distinction géographique entre cordillère de la Côte et précordillère a été incluse afin d'accorder plus d'importance à cette dernière.

Nous avons procédé à une analyse de la distribution des changements (1986-1999) et des types d'occupation du sol pour 1986 et 1999 en fonction de chaque variable. Chacune des variables (sauf os) a été mise en relation avec la carte des changements 1986-1999 afin d'étudier l'intensité de leur liaison grâce au coefficient de corrélation linéaire de Pearson, et ainsi estimer leur pouvoir explicatif. Les valeurs comprises entre -1 et 1 nous donnent des informations sur l'existence d'une relation linéaire entre la variable et les changements observés. Les coefficients de corrélation proches de 0 n'ont pas été systématiquement exclus dans la mesure où des relations non linéaires peuvent exister.

Les variables sous LCM sont de deux types: dynamiques et statiques. Les variables statiques traduisent l'aptitude pour chaque transition prise en considération et restent stable au cours du temps. Les variables dynamiques (proximité aux routes, proximité aux types d'occupation du sol déjà existant, infrastructures, etc.) sont recalculées à chaque pas durant le processus de prédiction. Le coefficient V de Cramer calcule la corrélation entre des variables et varie entre 0 et 1. Plus la corrélation est forte et plus le coefficient se rapprochera de 1 et vice versa. Selon Eastman (2004) la variable sera considérée comme acceptable si son coefficient est égal ou supérieur à 0.4, mais devra être rejetée si le coefficient est inférieur à 0.15. Malgré tout, il faut garder à l'esprit qu'une forte corrélation ne tient pas compte de la complexité des relations entre les variables.

Figure 3 : Principales variables explicatives



native (exploitation). Ces zones ont été laissées en tant que "forêt native" afin de faire émerger des zones potentielles de colonisation, en vue de leur conservation dans le futur.



# Validation des modèles

Afin de dépasser la subjectivité et le manque de précision d'une approche comparative purement visuelle entre une carte de référence (2008) et la simulation (2008), Pontius et al. (2004) montrent l'intérêt d'une comparaison statistique entre ces cartes. Tenant compte des récentes critiques portant sur les limites des indices de Kappa (Kno, Klocation, Kquantity, Kstandard) (Pontius et Millones, 2008), la validation se base sur la récente étude de Chen et Pontius (2010) qui propose une méthode permettant de quantifier et visualiser les relations entre les changements d'occupation du sol et des variables explicatives<sup>7</sup>. Notre intérêt s'est porté cependant sur la "budgétisation" des erreurs et des prédictions correctes. Les auteurs distinguent quatre catégories de pixels avec (i) les pixels corrects en raison d'une constance observée et prédite (*null successes [N]*), (ii) les erreurs en raison d'une constance observée mais prédite comme changée (*false alarms [F]*), (iii) les pixels corrects dus à un changement observé et prédit (*hits [H]*) et (iv) les erreurs dues à un changement observé mais prédit comme constant (*misses [M]*).

Pour juger la précision de la prédiction globale des changements à travers l'ensemble du paysage ils avancent une méthode permettant de mesurer les erreurs (en % du paysage) en raison de la quantité et celles en raison de l'allocation en se basant sur les résultats de la budgétisation susmentionnée (tableau 1).

Tableau 1 : Mesure des erreurs et des exactitudes

$$\begin{aligned} &OC : \text{changement observé en pourcentage du paysage} \\ &PC : \text{changement prédit} \\ &Q : \text{erreur en raison de la quantité des changements prédits} \\ &Q = |PC - OC| = |(F+H) - (M+H)| = |F-M| \\ &A : \text{erreur en raison de l'allocation des changements prédits} \\ &A = (F+M) - Q = 2 \times \text{MIN}(F, M) \\ &T : \text{erreur totale} \\ &T = F+M = Q + A \end{aligned}$$

L'erreur due à la quantité de changement prédite mesure le pourcentage d'imperfection de la correspondance entre la quantité de changement l'observée et prédite. L'erreur due à l'allocation mesure le degré d'approximation de la correspondance dans l'allocation spatiale des changements, compte tenu de la spécification de la quantité des changements dans les cartes de changements observés et prédits.

Afin de pouvoir appliquer cette méthode, la carte des changements observés entre 1999 et 2008 et celle des changements prédits doivent être croisées. Pour y parvenir la fonction CROSSTAB a été utilisée générant un tableau de contingence. Cette technique sera mobilisée dans un dernier temps afin de comparer et d'estimer la qualité des deux modèles, en se focalisant sur la prédiction des plantations issue de CA-Markov et celle de LCM. La carte met en exergue quatre états : (i) les zones non prédites comme plantations, (ii) les zones prédites comme plantations par les deux modèles, (iii) les plantations exclusivement prédites par CA-Markov et (iv) les plantations uniquement prédites par LCM.

## Résultats

La matrice de Markov entre 1999 et 2008 (tableau 2) indique la probabilité en pourcentage de chaque catégorie en 1999 de changer ou non pour une autre catégorie en 2008. On remarque une stabilité globale particulièrement marquée pour les plantations et l'eau. Les transitions s'opèrent de la catégorie "autre" vers FN et les plantations, et de FN vers la catégorie "autre" et plantations. Le tableau suppose qu'il y a une plus forte probabilité pour la FN de devenir "autre" que d'être convertie en monoculture (9% contre 1%). Cette dynamique suggère deux logiques de substitution. La première que l'on peut qualifier d'indirecte où la forêt est dégradée puis convertie en terre agricole ou simplement laissée à l'abandon, pour à terme faire l'objet d'une plantation. La deuxième qui est une substitution directe traduit un remplacement immédiat par du pin ou de l'eucalyptus. Les zones classées comme "autre" sont recolonisées par de la forêt native (14%) et seulement 4% par des plantations.

<sup>7</sup> L'objectif des auteurs n'est pas de comparer une image de référence à une simulation à une date t2, mais bien de comparer les changements observés et simulés entre une date t1 et t2.



Tableau 2 : Probabilité de transition (en %) entre les types d'occupation du sol

1999	2008			
	Plantation	Forêt native	Autre	Eau
Plantation	1.00	0.00	0.00	0.00
Forêt native	0.01	0.90	0.09	0.00
Autre	0.04	0.14	0.82	0.00
Eau	0.00	0.00	0.00	1.00

## Relations entre les changements observés et les variables explicatives

Le tableau 3 exprime le niveau d'association existant entre les principaux changements d'occupation du sol et les variables explicatives exposées précédemment. La distance aux villes semble ne pas jouer un rôle important pour les principales transitions observées entre 1986 et 1999. L'analyse de la distribution spatiale révèle cependant que la majorité d'entre-elles augmentent avec l'éloignement des centres urbains. La fréquence des changements d'occupation du sol s'amointrit à mesure que la distance aux routes, l'altitude et la pente augmentent, en témoignent les coefficients de corrélation linéaire supérieur à -0,50<sup>8</sup>. La proximité aux routes facilite l'accès et l'exploitation ressources forestières (forêt native et plantation). L'altitude représente également un frein à l'apparition de pins ou d'eucalyptus dans la mesure où les zones élevées, corrélées à de fortes pentes et à une raréfaction du réseau routier, rendrait toute plantation improductive. La déforestation (ForNat-Autre) s'explique elle aussi par la proximité aux routes, une déclivité faible et de basses altitudes (Lambin et Geist, 2007), de même que la proximité aux zones classées "autres" en 1986.

Les zones ayant transité de "autre" vers des plantations industrielles ont manifestement été influencées par la proximité aux parcelles de pins ou d'eucalyptus. Cette situation pouvant s'expliquer par la volonté des entreprises forestières d'accroître et de concentrer leur patrimoine forestier. La substitution des forêts natives par des monocultures est liée quant à elle à la proximité à "autre" en 1986. Cette dernière a un effet de polarisation dans la mesure où les plantations gagnées sur les forêts natives peuvent s'étendre plus facilement sur de nouveaux espaces dépourvus de végétation native.

Tableau 3 : Relation entre les transitions observées (1986-1999) et les variables spatiales. Les valeurs retenues sont en gras.

		DAC	DZU	ALT	PTE	DEN Plantation86	DEN Forêt native 86	DEN Autre 86
Trans. ForNat-Autre	Coef. Pearson	<b>-0.61</b>	0.11	<b>-0.58</b>	<b>-0.69</b>	0.10	-0.26	-0.41
	V de Cramer	0.09	0.04	0.10	0.07	0.05	<b>0.19</b>	<b>0.37</b>
Trans. ForNat-Plt	Coef. Pearson	<b>-0.66</b>	0.11	<b>-0.63</b>	<b>-0.72</b>	-0.14	-0.26	-0.42
	V de Cramer	0.05	0.06	0.09	0.04	<b>0.17</b>	0.06	0.13
Trans. Autre-Plt	Coef. Pearson	<b>-0.58</b>	0.11	<b>-0.65</b>	<b>-0.60</b>	<b>-0.86</b>	<b>-0.71</b>	-0.15
	V de Cramer	0.09	0.06	0.11	0.10	<b>0.17</b>	<b>0.21</b>	<b>0.20</b>
Trans. Autre-ForNat	Coef. Pearson	<b>-0.58</b>	0.12	<b>-0.67</b>	<b>-0.65</b>	-0.06	<b>-0.49</b>	-0.15
	V de Cramer	<b>0.15</b>	0.07	<b>0.16</b>	0.11	<b>0.19</b>	<b>0.41</b>	<b>0.37</b>

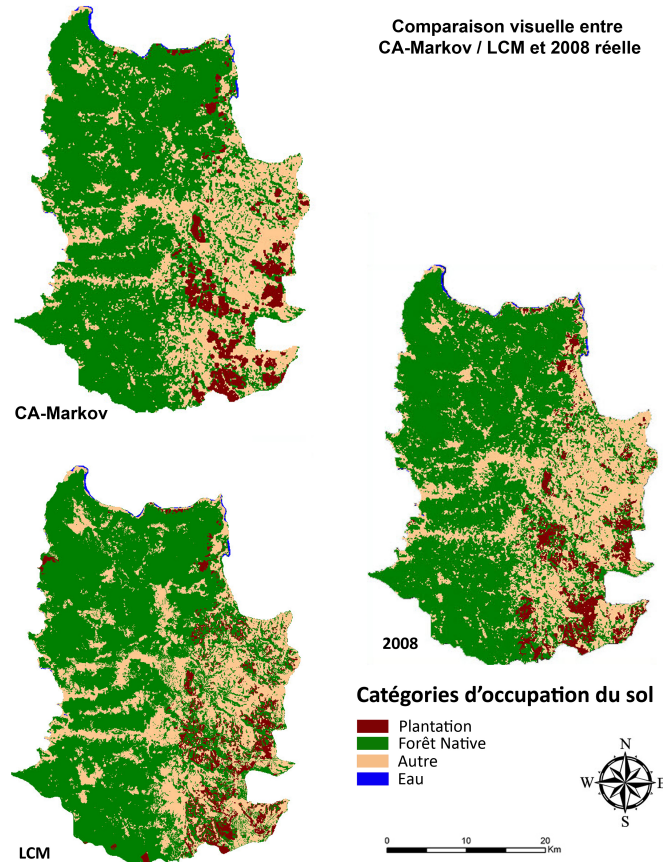
## Comparaison visuelle et validation des modèles

La figure 4 montre, pour 2008, l'occupation du sol observée et simulée. On remarque que la distribution des LUC a été correctement simulée, avec les forêts natives sur la Cordillère de la Côte (à l'ouest), les plantations sur la précordillère (à l'est) et la catégorie "autre" répartie sur l'ensemble de commune. CA-Markov et LCM sous-estiment l'expansion des monocultures, principalement dans le sud-est de la zone d'étude. L'analyse des concordances et des discordances entre les deux modèles et la carte de référence sera abordée en fin d'article.

Si l'approche visuelle apporte des premières pistes de réflexion, son caractère subjectif nous oblige à évaluer et valider statistiquement nos modèles.

<sup>8</sup> Seuls les coefficients de corrélation compris entre -1 et -0,5 (et 0,5 et 1) sont retenus.

Figure 4 : Occupation du sol en 2008 observée et simulée par CA-Markov et par LCM



## CA-Markov - Validation

La validation du modèle se déroule en deux étapes, en comparant d'abord l'analyse des changements prédits et observés entre 1999 et 2008, puis en se focalisant sur la catégorie des plantations.

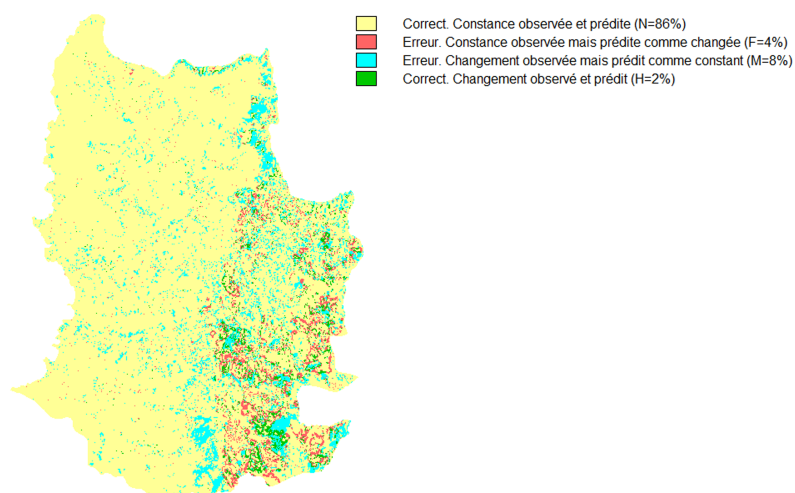
86% de la persistance observée (sur un total de 90%) entre 1999 et 2008 a été correctement prédite (N) (tableau 4 et figure 5). La prédominance de la forêt native et sa stabilité à l'échelle communale comparativement aux autres types d'occupation explique en partie ce résultat. Les erreurs en raison d'une constance observée mais prédite comme changée sont de 4% (F), alors que les *misses* qui traduisent le processus inverse atteignent les 8%. Enfin les changements observés ayant été correctement prédits par le modèle sont de 2% (H). Le total des changements observés ( $OC = M + H$ ) est de 10% alors que les changements totaux prédits ( $PC = F + H$ ) ont été sous-estimés avec 6%. La précision de la prédiction globale des changements à travers l'ensemble du paysage est la suivante :

Erreur de quantité ( $Q$ ) =  $|F - M| = 4\%$     Erreur d'allocation ( $A$ ) =  $(F + M) - Q = 12\%$     Erreur totale ( $Q + A$ ) =  $14\%$

Tableau 4 : Matrice de contingence entre les changements observés et les changements prédits (CA-Markov)

%		Changements observés 99-08		
		0	1	Total
Changements prédits 99-08	0	0.86	0.08	0.94
	1	0.04	0.02	0.06
	Total	0.90	0.10	1.00

Figure 5 : Exactitudes et erreurs de la prédiction – CA-Markov



Les taux d'exactitudes et d'erreurs renseignent tous les LUC confondus, mais ne donnent pas de précision pour chacune d'elles. Pour obtenir ces résultats, et à l'instar de la figure 5, un tableau de contingence à trois entrées a été calculée, avec en ligne la carte de 2008, en colonne la carte de 1999 et en troisième entrée la simulation de 2008. La validation des changements par type d'occupation du sol montre que le modèle prédit 40% de changement (PC) pour les plantations alors que les changements réels (OC) pour cette catégorie n'ont été que de 18%. Les erreurs dues à la quantité sont de 22% alors que celles dues à la distribution sont nulles.

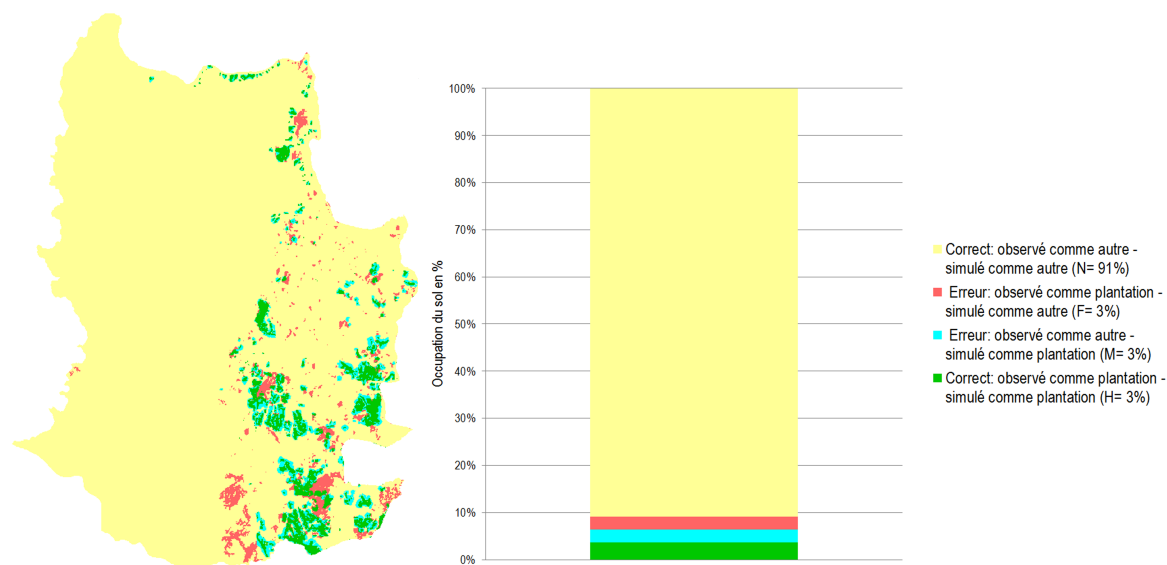
La confrontation simulée – observée (tableau 5) exprime la concordance et la discordance spatiale des plantations. La valeur 1 représente les plantations alors que le 0 caractérise un autre type d'occupation du sol (forêt native ou "autre"). En excluant la part des autres LUC, on remarque que 57% des plantations ont été correctement prédites par le modèle, et que par conséquent ce dernier a omis près de 43% d'entre-elles (4207,32 ha).

Les figures 6 et 7 traduisent cartographiquement et graphiquement le pourcentage d'exactitudes et d'erreurs pour les plantations. Le segment supérieur (figure 7) exprime un taux d'exactitude (N) lié à un type d'occupation du sol (autre que plantation) observé et simulé en 2008. Le résultat extrêmement élevé (91%) s'explique par la faible part qu'occupent les plantations à l'échelle de la commune comparativement aux autres types d'occupation du sol. Le segment suivant correspond à une première erreur due à l'observation de plantations en 2008 n'ayant pas été simulées (F=3%). Le même résultat est obtenu *a contrario* pour les plantations simulées mais observées comme autre (M=3%). Enfin, 4% du paysage a été simulé et observés comme étant des plantations (H).

Tableau 5 : Concordance et discordance spatiale des plantations (CA-Markov)

Hectares		plantation 08 observée		
		non plantation	plantation	Total
plan tatio n CA- Mar kov	non plantation	0.91	0.03	0.94
	plantation	0.03	0.04	0.06
	Total	0.94	0.06	1.00

Figure 6 et 7 : Exactitudes et erreurs de la prédiction pour les plantations, entre 2008 réel et 2008 simulée – CA-Markov



## Analyse des résidus (CA-Markov)

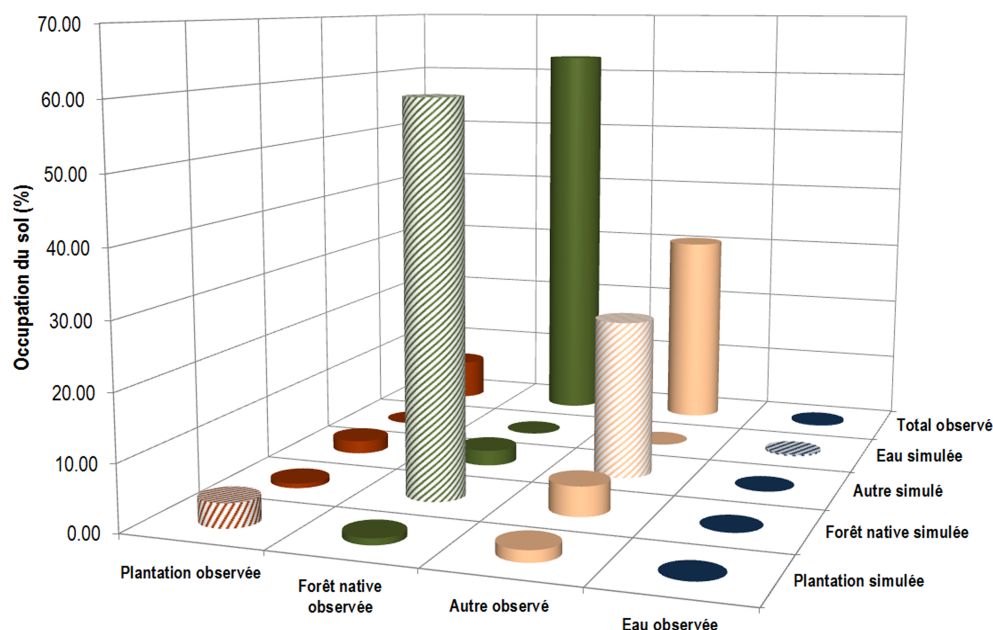
La comparaison entre l'occupation du sol observée et simulée (tableau 6) met en exergue deux informations : (i) la part des concordances spatiales (en gris) et (ii) les résidus qui représentent les types d'occupation du sol simulés qui ne correspondent pas à l'observation. Ces informations sont également décrites par la figure 8 où les concordances sont traduites par des cylindres à rayures.

Sur les 6,4% de plantations simulées en 2008, 3,7% concordent avec l'observation. Les 2,73% restant (0,99% de forêt native et 1,74% de "autre") correspondent aux résidus (ou M). Ils se localisent, pour la plupart, à proximité des plantations observées (figure 6). *A contrario*, les autres résidus (F) (2,02% de "autre" et 0,71% de forêt native) traduisent des erreurs de simulation où le modèle omet l'apparition de nouvelles plantations.

Tableau 6 : Comparaison entre l'occupation du sol observée et simulée (CA-Markov)

	Plantation	Forêt native	Autre	Eau	Total 2008 simulé
Plantation	3.67	0.99	1.74	0.00	6.40
Forêt native	0.71	59.33	4.52	0.00	64.56
Autre	2.02	2.38	24.38	0.00	28.78
Eau	0.00	0.00	0.00	0.26	0.26
Total 2008 observé	6.40	62.71	30.64	0.26	100

Figure 8 : Observation des concordances (cylindres à rayures) et des résidus (autres cylindres) entre l'occupation du sol observée et simulée (CA-Markov) (en pourcentage).



## LCM – Validation

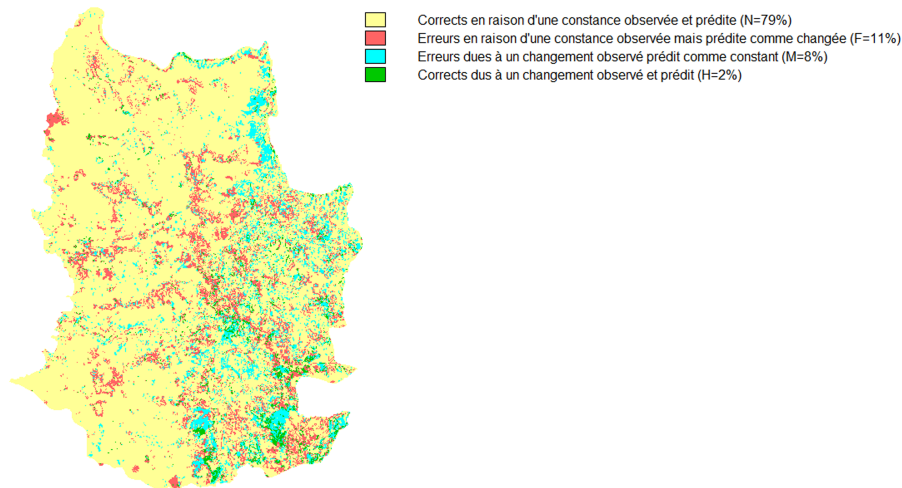
79% de la constance observée entre 1999 et 2008 a été correctement prédite (N) (tableau 7 et figure 9). L'explication est identique à celle exposé précédemment. Les erreurs en raison d'une constance observée mais prédite comme changée sont de 11% (F), alors que les *misses* qui traduisent le processus inverse atteignent les 8%. Enfin les changements observés ayant été correctement prédits par le modèle est de 2% (H). Le total des changements observés ( $OC = M + H$ ) est de 10% alors que les changements totaux prédits ( $PC = F + H$ ) ont été sous-estimés avec 8%. La précision de la prédiction globale des changements à travers l'ensemble du paysage est la suivante :

Erreur de quantité ( $Q$ ) =  $|F - M| = 3\%$     Erreur d'allocation ( $A$ ) =  $(F + M) - Q = 16\%$     Erreur totale ( $Q + A$ ) =  $19\%$

Tableau 7: Matrice de contingence entre les changements observés et les changements prédits (LCM)

%		Changements observés 99-08		
		0	1	Total
Changements prédits 99-08	0	0.79	0.08	0.87
	1	0.11	0.02	0.13
	Total	0.90	0.10	1.00

Figure 9 : Exactitudes et erreurs de la prédiction – LCM



Les résultats OC et PC diffèrent de CA-Markov car leur calcul se base sur les résultats de la prédiction. La validation des changements par type d'occupation du sol montre que le modèle prédit 58% de changement (PC) pour les plantations alors que les changements réels (OC) pour cette catégorie n'ont été que de 13%<sup>9</sup>. Les erreurs dues à la quantité sont de 45% alors que celles dues à la distribution sont nulles.

De la confrontation entre les plantations réellement observées et prédites (tableau 8) découle une concordance spatiale de 3%. En excluant la part des autres LUC, cette concordance atteint les 50%. En effet, LCM a simulé autant de plantations qu'elle en a omises (4903,02 ha contre 4903,20 ha).

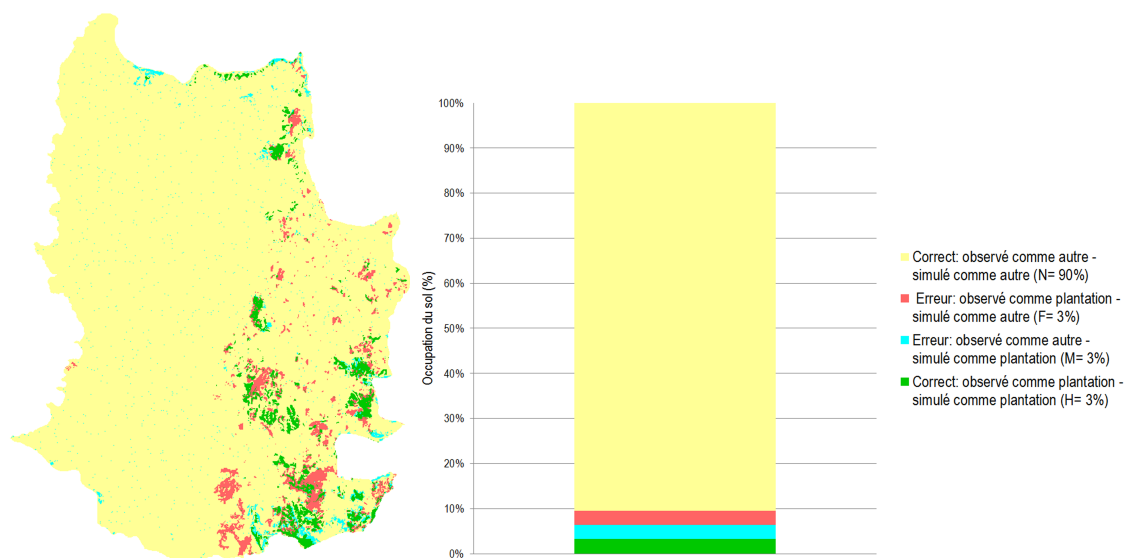
Les erreurs remarquables dans les figures 10 et 11 correspondent à la part des plantations observées mais n'ayant pas été simulées en tant que telles (F = 3%), et la part des plantations simulées mais observées comme autres (M = 3%). Le taux d'exactitude pour les plantations observées et celles prédites atteint également les 3% (H sur le graphique). Enfin la plus forte concordance (90%) exprime les autres catégories d'occupation du sol, d'où le taux élevé d'exactitude (N sur le graphique).

Tableau 8 : Concordance et discordance spatiale des plantations (LCM)

Hectares		plantation 08 observée		
		non plantation	plantation	Total
plantation LCM	non plantation	0.90	0.03	0.94
	plantation	0.03	0.03	0.06
	Total	0.94	0.06	1.00

<sup>9</sup> Les changements observés (OC) diffèrent des résultats exposés pour CA-Markov (18%) ce qui pourrait sembler erroné. Ceci s'explique par le calcul du OC qui tient compte des résultats de la prédiction (H + M) et que ces derniers diffèrent entre les deux simulations.

Figure 10 et 11 : Exactitudes et erreurs de la prédiction pour les plantations, entre 2008 réel et 2008 simulée – LCM



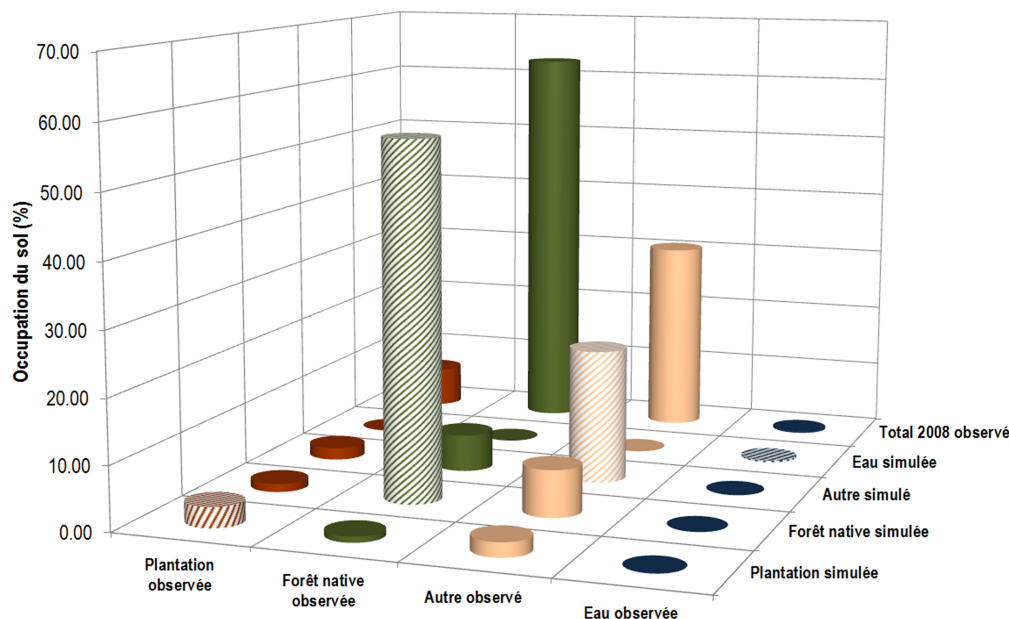
## Analyse des résidus (LCM)

Comme pour l'analyse des résidus (CA-Markov), la part des concordances spatiales est surlignée en gris et le reste correspond aux résidus. Sur les 6,4% de plantations simulées en 2008, 3,2% concordent avec l'observation (tableau 9). Les 3,2% restant (0,92% de forêt native et 2,26% de "autre") correspondent aux résidus (ou M). Contrairement à CA-Markov, ils se répartissent de manière diffuse et fragmentée sur le piémont sans aucune relation de proximité avec des plantations existantes, et apparaissent plus homogènes au sud et à l'ouest de la commune (figure 10). *A contrario*, les autres résidus (F) (1,90% de "autre" et 1,28% de forêt native) traduisent des erreurs de simulation où le modèle omet l'apparition de nouvelles plantations.

Tableau 9 : Comparaison entre l'occupation du sol observée et simulée (LCM)

	Plantation	Forêt native	Autre	Eau	Total 2008 simulé
Plantation	3.22	0.92	2.26	0.00	6.40
Forêt native	1.28	55.94	7.34	0.00	64.56
Autre	1.90	5.85	21.04	0.00	28.78
Eau	0.00	0.00	0.00	0.26	0.26
Total 2008 observé	6.40	62.71	30.64	0.26	100

Figure 12 : Observation des concordances (cylindres à rayures) et des résidus (autres cylindres) entre l'occupation du sol observée et simulée (LCM) (en pourcentage).

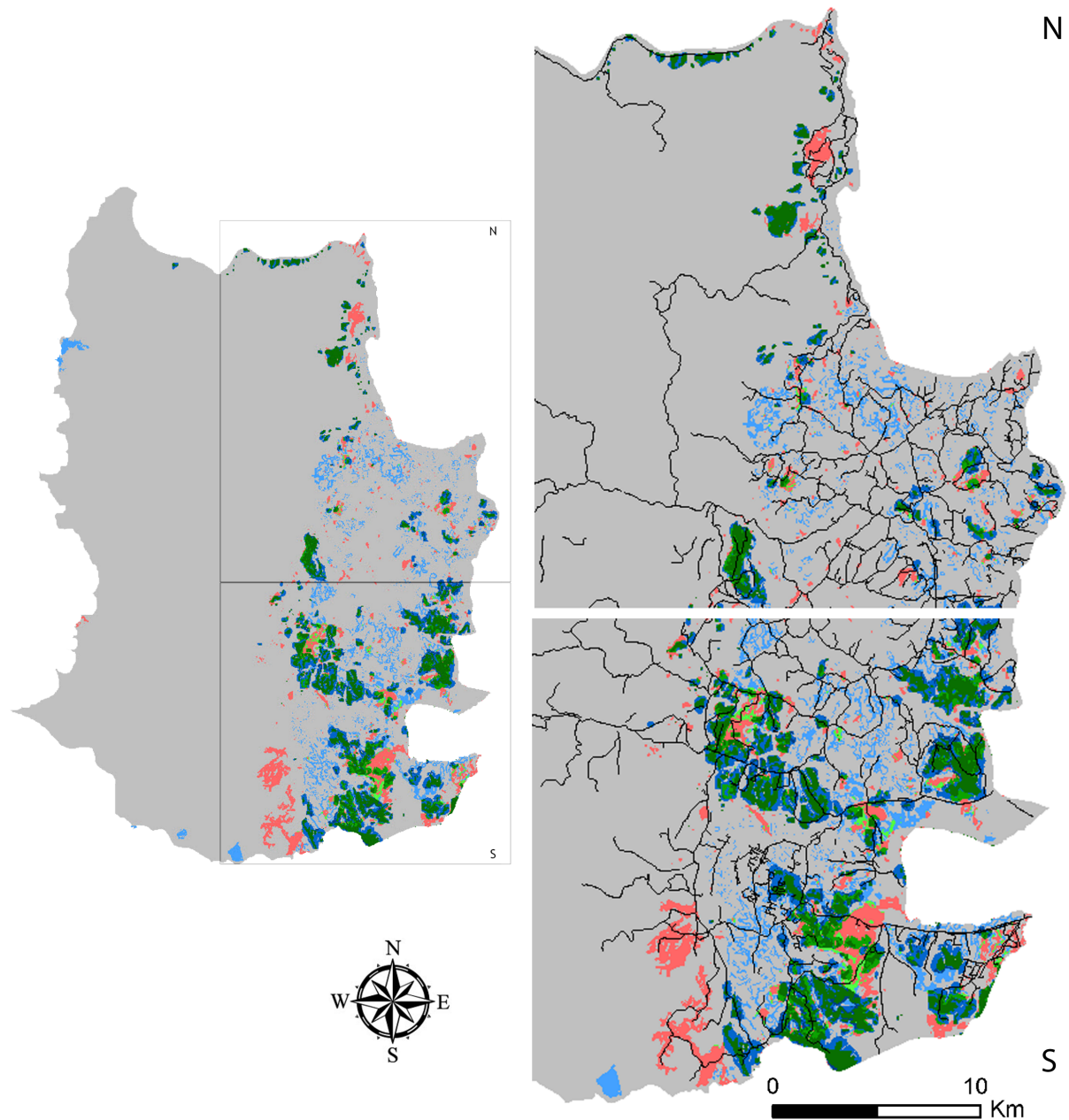


## Des espaces à enjeux

La figure 13 exprime la congruence entre plantations observées et simulées par les deux modèles ainsi que chacun des modèles pour 2008. Les deux modèles prédisent les plantations réellement observées de la même manière – 4 504,7 hectares, soit 45,6%. Pris isolément, CA-Markov génère des résultats plus proche de la réalité que LCM avec 5 672,3 ha (57,4 %) contre 4 976,5 ha (50,4 %) (figure 14). *A contrario*, les erreurs induites par la simulation de "non plantation" par les deux modèles mais observées comme plantation s'élèvent à 3 735,45 ha (en rose sur la figure 13). Les erreurs communes aux deux modèles se chiffrent à 639,5 ha, alors qu'individuellement CA-Markov compte 4 212,5 ha de surestimation et LCM 4 903,2 ha (en dégradé de bleu). Ces erreurs apportent une dimension non négligeable quant à l'évolution future des monocultures. Bien que ces zones ne soient pas réellement observables elles représentent des aires potentielles pour de nouvelles plantations. Si la plupart d'entre-elles se situent sur la précordillère, on remarque que certaines (LCM) occupent des zones totalement isolées mais prédites malgré tout en tenant compte du panel de variables explicatives.

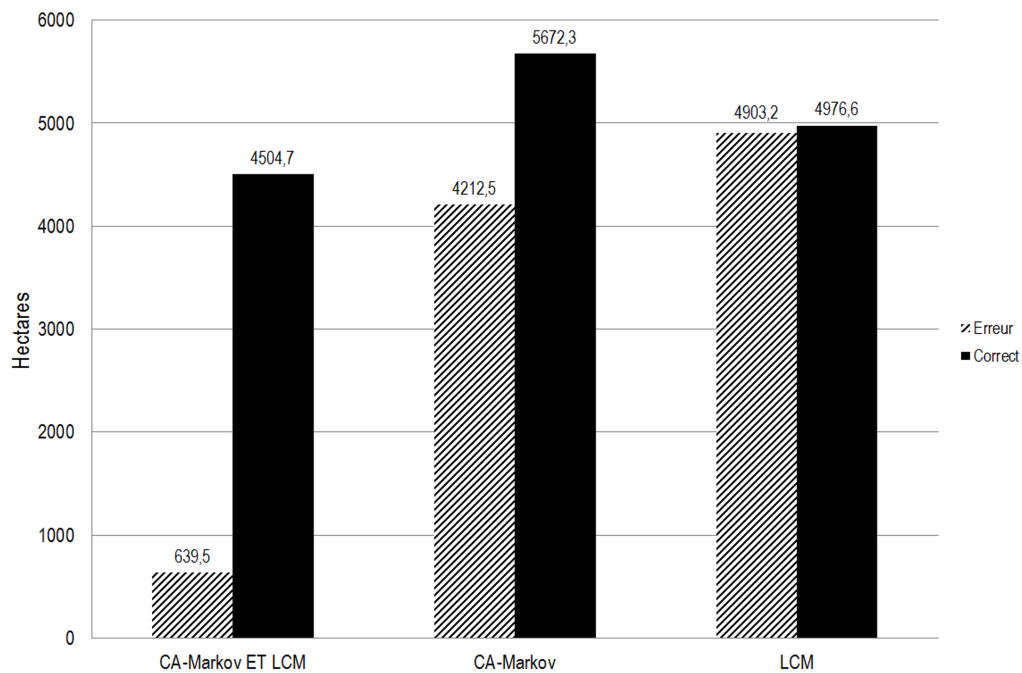


Figure 13: Comparaison des plantations simulées par CA-Markov et LCM avec les plantations observées en 2008. Carte globale de la commune de San Juan de la Costa et deux zooms sur la précordillère



- Correct: simulé plantation (CA-Markov + LCM) - 45,6 %
- Correct: simulé plantation CA-Markov - 57,4 %
- Correct: simulé plantation LCM - 50,4 %
- Correct: simulé "non plantation" (CA-Markov + LCM)
- Erreur: simulé "non plantation" (CA-Markov + LCM) - observé plantation
- Erreur: simulé plantation (CA-Markov + LCM) - observé "non plantation" - 6,5 %
- Erreur: simulé plantation CA-Markov - observé "non plantation" - 42,6 %
- Erreur: simulé plantation LCM - observé "non plantation" - 49,6 %
- ~ Réseau routier

Figure 14 : Erreurs et exactitudes dans la prédiction des plantations (CA-Markov et LCM) comparées aux plantations observées



## Discussion et conclusion

Malgré des quantités proches de la réalité, les modèles présentent, dans l'ensemble, des difficultés à prédire les changements entre 1999 et 2008 ; toutefois ils confirment les tendances d'évolution. En effet, l'analyse des changements, aussi bien à l'échelle du paysage qu'au niveau des plantations prises isolément, met en exergue des résultats intéressants mais qu'il convient de nuancer. Ceci est imputable à la part importante de la constance observée et simulée, 86% pour CA-Markov et 79% pour LCM à l'échelle du paysage.

Néanmoins, la comparaison des plantations observées et simulées en 2008 donnent des résultats acceptables quant à l'allocation spatiale. Si LCM prédit autant de plantations qu'il en omet, CA-Markov affiche des résultats plus proches de l'observé. En excluant les zones observées comme "autre" et simulées comme "autres" qui "écrasent", une fois de plus, les résultats par leur surface, les deux modèles affichent des pourcentages de -prédictions correctes- équivalentes aux erreurs (sauf pour CA -Markov dont les pourcentages d'exactitude sont supérieurs aux erreurs). L'observation des erreurs (observé comme autre – simulé comme plantation) révèle que celles-ci se localisent à proximité des plantations réellement observées pour CA-Markov, et de façon aléatoire pour LCM. A l'échelle du paysage les erreurs d'allocations spatiales issues de LCM ont été nettement supérieures à celles de CA-Markov.

L'analyse comparative des deux modèles avec la simulation observée indique que les plantations prédites par CA-Markov ont un taux de concordance plus élevé que celles simulées par LCM. La simulation des catégories d'occupation du sol par une approche supervisée (CA-Markov) permet de générer des cartes plus réalistes. Les cartes d'aptitude générées par le modélisateur optimisent la probabilité des changements contrairement à LCM qui se base sur les probabilités de transitions. Ce dernier fait cependant émerger des zones potentielles de plantations qui ne seraient pas apparues si l'on avait privilégié exclusivement une approche experte. Une critique commune aux deux modèles peut être avancée dans la mesure où l'établissement de nouvelles plantations, de même que les ruptures n'aient pas été anticipées.

La limite des modèles se trouve principalement dans la disponibilité des données mobilisées. Malgré tout certains auteurs ont montré que l'ajout de données supplémentaires pouvait être limité par leur nature non quantifiable et leur indisponibilité en format digital (Schneider et Pontius, 2001). La complexité du système forestier dans la commune de San Juan de la Costa ne peut effectivement se résumer à un nombre restreint de facteurs. Aux variables utilisées s'ajoutent des facteurs explicatifs dont la spatialisation est difficile à mettre en œuvre voire impossible. Nos entretiens ont permis d'identifier des facteurs politiques et institutionnels

(pauvreté, chômage, conflits, coût de la terre, etc.), des facteurs démographiques (migration, distribution de la population), des facteurs culturels (consommation des ménages) et des facteurs économiques (coût du travail et du capital).

La question qui se pose alors est de savoir à quel degré de complexité (induit par la quantité de variables incluses dans les modèles) le modélisateur et le modèle se limitent-ils. De même que la prise en compte de la totalité hypothétique des variables explicatives permettrait-elle une précision optimale de la prédiction ? L'importance avérée du régime foncier sur la dynamique des plantations nous amène à accorder une part plus conséquente aux propriétés des entreprises forestières. Des données foncières digitalisées seront intégrées à notre modèle afin d'améliorer l'allocation spatiale des monocultures.

La validation de ces modèles constitue une première étape à la modélisation prospective des plantations industrielles à l'horizon 2017. Bien que CA-Markov présente des résultats plus concluants que LCM et semble être à même de mieux prendre en compte les tendances d'évolution, ce dernier, de par son caractère non-supervisé (boîte noire), pourrait être plus adapté pour simuler des scénarii contrastés en restant totalement indépendant de la connaissance des experts. L'élaboration d'images issues de scénarii prospectifs tendanciels et contrastés favorisera l'identification de zones à enjeux socio-environnementaux concernant d'une part le milieu de vie des communautés indigènes mapuches et d'autre part la préservation des forêts natives.

## Remerciements

Les auteurs sont reconnaissant au projet ECOS-Sud Chili n° C07H03 intitulé "La forêt de la cordillère côtière continentale dans le sud du Chili : Dynamiques contemporaines et modélisations prospectives".

## Bibliographie

Altamirano A., Lara A., 2010, "Deforestación en ecosistemas templados de la precordillera andina del centro-sur de Chile", *Bosque*, vol. 31, No.1, 53-64.

Arana Silva M.T., 2009, "¿Hacia dónde va la industria chilena? Situación actual, inversiones y desafíos", Seminario para periodistas, Santiago, CORMA, 27 p.

Arroyo M.T.K., Marticorena C., Matthei O. et Cavieres L., 2000, "Plant Invasions in Chile: Present Patterns and Future Predictions", in Mooney H.A., Hobbs R.J. (eds.), *Invasive species in a changing world*, Washington, D.C., 385-421.

Balster H., 1998, "Cellular automata models for vegetation dynamics", *Ecological Modelling*, vol. 107, 113-125.

Balster H., 2000, "Markov chain models for vegetation dynamics", *Ecological Modelling*, vol. 126, 139-154.

Barrué-Pastor M., 2004, "Concept de propriété, conflits territoriaux et développement durable des forêts et des communautés Mapuche-Huilliches", in : Barrué-Pastor M. (dir.), *Forêts et développement durable au Chili. Indianité Mapuche et mondialisation*, Toulouse, Presse Universitaire du Mirail, 193-238.

Biblioteca del Congreso Nacional de Chile, 2008, "Reporte Estadístico Comunal. San Juan de la Costa", 24 p.

Bonnefoy J-L., Bousquet F., et Rouchier J., 2001, "Modélisation d'une interaction individus, espace et société par les systèmes multi-agents : pâture en forêt virtuelle", *L'Espace Géographique*, vol. 1, 13-25.

Bull G.Q., Bazett M., Schwab O., Nilson S., White A. et Maginnis S., 2006, "Industrial forest plantation subsidies: Impacts and implications", *Forest Policy and Economics*, vol. 9, 13-31.

Camacho Olmedo M.T., Paegelow M., et Martinez P.J., 2007, "Modélisation géomatique rétrospective des paysages par évaluations multicritères et multiobjectifs", *Cybergeo : European Journal of Geography*, Systèmes, Modélisation, Géostatistiques, <http://cybergeo.revues.org/4811>, No.365, 9 mars 2007, 24 p.

Cannell M.G.R., 1999, "Environmental impacts of forest monocultures: water use, acidification, wildlife conservation, and carbon storage", *New Forests*, vol. 17, 239-262.

Cavelier J., Tecklin D., 2005, "Conservación de la Cordillera de la Costa: un desafío urgente en la Ecorregión Valdiviana", in Smith-Ramírez C., Armesto J.J., Valdovinos C. (eds.), *Historia, biodiversidad y ecología de los bosques costeros de Chile*, Edición Universitaria, Santiago de Chile, 632-644.

Chen H., Pontius Jr R.G., 2010, "Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable", *Landscape Ecology*, vol. 25, 1319-1331.

CONAF, CONAMA, et BIRF, Mars 1999, "Catastro y Evaluación de los Recursos Vegetacionales Nativos de Chile. Informe nacional con variables ambientales", Universidad Austral de Chile, Pontificia Universidad Católica de Chile, Universidad Católica de Temuco.

Collectif ComMod, 2005, "La modélisation comme outils d'accompagnement", *Natures Sciences Sociétés*, vol. 13, 165-168.

Colasanti R.L., 1991, "Discussions of the possible use of neural network algorithms in ecological modeling", *Binary: Computing in Microbiology*, vol. 3, No.1, 13-15.

Condorbes S., Durance P., 2004, *Attitudes prospectives. Eléments d'une histoire de la prospective en France après 1945*, Paris, L'Harmattan.

Coquillard P., Hill D.R.C., 1997, *Modélisation et simulation d'écosystèmes. Des modèles déterministes aux simulations à événements discrets*, Paris, Edition Masson.

Costanza R., Voinov A., 2004, *Landscape simulation modeling. A spatially explicit, dynamic approach*, New York, Springer.

Daudé E., 2005, "Systèmes multi-agents pour la simulation en géographie: vers une Géographie Artificielle", in : Germond Y. (dir.), *Modélisations en géographie. Déterminismes et complexités*, Paris, Lavoisier, 353-380.

De Jouvenel H., 1999, "La démarche prospective. Un bref guide méthodologique", *Futuribles*, No.247, 24 p.

Donoso P.J., Otero L.A., 2005, "Hacia una definición de país forestal: ¿Dónde se sitúa Chile?", *Bosque*, vol. 26, No.3, 5-18.

Eastman J.R., 2006, *IDRISI Andes. Guide to GIS and Image Processing*, Worcester.

Eastman J.R., 2009, "The Land Change Modeler for Ecological Sustainability", *IDRISI Andes. Guide to GIS and Image Processing*, Worcester, 234-256.

Echeverria C., Coomes D., Salas J., Rey-Benayas J-M., Lara A. et Newton A., 2006, "Rapid deforestation and fragmentation of Chilean Temperate Forests", *Biological Conservation*, 481-494.

Echeverria C., Coomes D., Hall M. et Newton A.C., 2008, "Spatially explicit models to analyze forest loss and fragmentation between 1976 and 2020 in southern Chile", *Ecological modeling*, vol. 212, 439-449.

Evans J., Turnbull J.W., 2004, *Plantation forestry in the tropics: the role, silviculture, and use of planted forests for industrial, social, environmental, and agroforestry purposes*, New-York, Oxford University Press.

Ferber J., 2006, "Concepts et méthodologies multi-agents", in : Amblard F., Phan D. (dir.), *Modélisation et Simulation Multi-Agents. Application pour les Sciences de l'Homme et de la Société*, Londres, Hermès science & Lavoisier, 23-27.

Glenn J.C., Gordon T.J., et Florescu E., 2010, *2010 State of the Future*, The Millennium Project.

Godet M., 1986, "Introduction à la Prospective. Seven key ideas and one scenario method", *Futures*, 134-157.

Hatem F., 1993, *La Prospective. Pratiques et Méthodes*, Paris, Economica.

Houet T., 2006, "Occupation des sols et gestion de l'eau: Modélisation prospective en paysage agricole fragmenté (Application au SAGE Blavet) ", Thèse de doctorat, Université de Rennes 2 - Haute Bretagne.

Houet T., 2006, "Modélisation prospective de l'occupation du sol en zone agricole intensive dans la France de l'Ouest", *Norois*, <http://norois.revues.org/pdf/2025>, No.198, 2006/1, 19 p.

Houet T., Hubert-Moy L. et Tissot C., 2008, "Modélisation prospective spatialisée à l'échelle locale : approche méthodologique. Application à la gestion de l'eau en Bretagne", *Revue internationale de Géomatique*, vol. 18, No.3, 345-373.

Hubert-Moy L., Corgne S., Houet T. et Tissot C., 2006, "Modélisation prédictive et prospective des changements de l'occupation des sols", *Conférence Interactions Nature-Sociétés : Analyse et modèles*, 3-6 mai 2006, La Baule.

INstituto FORestal (INFOR), 2008, *Anuario Forestal 2008*, Santiago de Chile, Centro de Información Forestal (CIF).

Irwin E.G., Geoghegan J., 2001, "Theory, data, methods: developing spatially explicit economic models of land use change", *Agriculture, Ecosystems and Environment*, vol. 85, 7-23.

Karul, C. and S. Soyupak, 2006, "A comparison between neural network based and multiple regression models for Chlorophyll-a estimation", in Recknagel, F. (ed), *Ecological Informatics*, Berlin, Springer.

Lambin E.F., Geist H. J., Ellis E. (topic editor), 2007, "Causes of land-use and land-cover change", *Encyclopedia of Earth*, Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment, 5 p.

Lara A., 1996, "Una propuesta general de silvicultura para Chile", *Ambiente y Desarrollo*, vol. 12, No.91, 31-40.

Lek S., Guégan J.F., 1999, "Artificial neural networks as a tool in ecological modelling, an introduction", *Ecological Modelling*, vol. 120, 65-73.

Mas J.F., Puig H., Palacio J.L. et Sosa-López A., 2004, "Modelling deforestation using GIS and artificial neural networks." *Environmental Modelling & Software*, vol. 19, 461-471.

Mas J.F., Kolb M., Houet T., Paegelow M. et Camacho Olmedo M.T.,(sous presse), "Eclairer le choix des outils de simulation des changements des modes d'occupation et d'usages des sols. Une approche comparative", *Revue Internationale de Géomatique*.

Medina Vásquez J., Ortégón E., 2006, *Manual de prospectiva y decisión estratégica: bases teóricas e instrumentos para América Latina y el Caribe*, Santiago de Chile, CEPAL.

Morán-Ordóñez A., Suárez-Seoane S., Calvo L. et De Luis E., 2011, "Using predictive models as a spatially explicit support tool for managing cultural landscapes", *Applied Geography*, vol. 31, 839-848.

Muñoz C.P., 2004, "Alerce: El Padre Forestal de Chile." *Chile Forestal*, 30-32.

Myers N., Mittermeier R.A., Mittermeier C.G., Da Fonseca G.A.B. et Kent J., 2000, "Biodiversity hotspots for conservation priorities", *Nature*, vol. 403, No.24, 853-858.

Oñate-Valdivieso F., Bosque Sendra J., 2010, "Application of GIS and remote sensing techniques in generation of land use scenarios for hydrological modeling", *Journal of Hydrology*, vol. 395, 256-263.

Paegelow M., Villa N., Cornez L., Ferraty F., Ferré L. et Sarda P., 2004, "Modélisations prospectives de l'occupation du sol. Le cas d'une montagne méditerranéenne ", *Cybergeo : European Journal of Geography*, Systèmes, Modélisation, Géostatistiques, <http://cybergeo.revues.org/2811>, No.295, 6 décembre 2004, 18 p.

Paegelow M. and Camacho Olmedo M.T., 2005, "Possibilities and limits of prospective GIS land cover modelling—a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain)", *International Journal of Geographical Information Science*, vol. 19, No.6, 697-722.

Parizeau M., 2004, "Le perceptron multicouche et son algorithme de rétropropagation des erreurs", Laval, Département de génie électrique et de génie informatique.

Parker D.C., Manson S.M., Janssen M.A., Hoffman M.J. et Deadman P., 2003, "Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review", *Annals of the Association of American Geographers*, vol. 93, No.2, 314-337.

Perry G.L.W., Millington J.D.A., 2008, "Spatial modelling of succession-disturbance dynamics in forest ecosystems: Concepts and examples", *Perspectives in Plant Ecology, Evolution and Systematics*, vol. 9, 191-210.

Pontius Jr R.G., Huffaker D. et Denman K., 2004, "Useful techniques of validation for spatially explicit land-change models", *Ecological Modelling*, vol. 179, 445-461.

Pontius Jr R.G., Malanson J., 2005, "Comparison of the structure and accuracy of two land change models", *International Journal of Geographical Information Science*, vol. 19, No.2, 243-265.

Pontius Jr R.G., Millones M., 2008, "Problems and solutions for kappa-based indices of agreement", *International Conference*, Mytilene, Lesvos, Greece.

Saaty T., 1977, "A scaling method for priorities in hierarchical structures", *Journal of mathematical psychology* vol. 15, No. 3, 234-281.

Saez Villalobos N., 2002, "Clasificación de espacios forestales nativos en el norte de la Isla de Chiloé", *Revista Geográfica de Chile Terra Australis*, 249-265.

Sánchez Álvarez J., 2004, "Politiques publiques, institutions et dynamiques forestières au sud du Chili" in: Barrué-Pastor (dir.), *Forêts et développement durable au Chili. Indianité mapuche et mondialisation*, Toulouse, Presse Universitaire du Mirail, 49-64.

Schneider L.C., Pontius Jr R.G., 2001, "Modeling land-use change in the Ipswich watershed, Massachusetts, USA", *Agriculture, Ecosystems & Environment*, vol. 85, 83-94.

Schulz J.J., Cayuela L., Echeverria C., Salas J. et Rey-Benayas J-M., 2010, "Monitoring land cover change of the dryland forest landscape of Central Chile (1975-2008)", *Applied Geography*, vol. 30, 436-447.

Soares-Filho B.S., Rodrigues H.O., et Costa W.L., 2009, *Modeling Environmental Dynamics with Dinamica EGO*, Belo Horizonte – Brazil

Veldkamp A., Fresco L.O., 1996, "CLUE: a conceptual model to study the Conversion of Land Use and its Effects", *Ecological Modelling*, vol. 85, 253-270.

Veldkamp A., Lambin E. F., 2001, "Predicting land-use change", *Agriculture, Ecosystems & Environment*, vol. 85, No.1-3, 1-6.

Veldkamp A., Verburg P., 2004, "Modelling land use change and environmental impact", *Journal of Environmental Management*, vol. 72, 1-3.

Verburg P., Soepboer W., Veldkamp A., Limpiada R., Espaldon V. et Mastura S.S.A., 2002, "Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model", *Environmental Management*, vol. 30, No.3, 391-405.

Verburg P., Schot P.P., Dijst M.J. et Veldkamp A., 2004, "Land use change modelling: current practice and research priorities", *Geo Journal*, vol. 61, 309-324.

Villagran C., 1991, "Historia de los bosques templados del sur de Chile durante el Tardiglacial y Postglacial", *Revista Chilena de Historia Natural*, vol. 64, 447-460.

Vitousek P.M., Mooney H. A., Lubchenco J. et Melillo J-M., 1997, "Human Domination of Earth's Ecosystems", *Science*, No. 277, 494-499.

Voiron-Canicio C., 2006, "L'espace dans la modélisation des interactions nature-société", *Colloque Interactions Nature-Sociétés : Analyse et modèles*, 3-6 mai 2006, La Baule.

Winjum J.K., Schroeder P.E., 1997, "Forest plantations of the world: their extent, ecological attributes, and carbon storage", *Agricultural and Forest Meteorology*, vol. 84, 153-167.

Wulder M.A., White J.C., Goward S.N., Masek J.G., Irons J.R., Herold M., Cohen W.B., Loveland T.R., Woodcock C.E., 2008, "Landsat continuity: Issues and opportunities for land cover monitoring", *Remote Sensing of Environment*, vol. 112, 955-969.

Wyman M.S., Stein T.V., 2010, "Modelling social and land-use/land-cover change data to assess drivers of smallholder deforestation in Belize", *Applied Geography*, vol. 30, 329-342.